BBC RD 1984/14

RESEARCH DEPARTMENT REPORT

Electronic generation of text for television

N.E. Tanton, M.A., C.Eng., M.I.E.E., M.Inst.P.

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION

November 1984

BBC RD 1984/14
UDC 621.397

ELECTRONIC GENERATION OF TEXT FOR TELEVISION

N.E. Tanton, M.A., C.Eng., M.I.E.E., M. Inst.P.

Summary

The image quality of present-day television character generators is compared with the quality of text images derived by scanning real artwork. To illustrate and identify the causes of image distortion (aliasing) prevalent in these and other synthetic television images, the spectral consequences of scanning a scene and of representing any image with a two-level (1-bit) signal are discussed.

A method of electronically generating and displaying filtered text images for television is described together with a comparison of various filtering methods. The resulting images are largely free from aliasing and have a quality equal to or better than that of scanned artwork.

Issued under the Authority of

Head of Research Department

Research Department, Engineering Division,
BRITISH BROADCASTING CORPORATION

November, 1984
(PH-262)

ELECTRONIC GENERATION OF TEXT FOR TELEVISION

Section Title Page

Summary .... Title Page

1. Introduction 1

2. The effects of scanning 2

2.1. Sampling theory 2
2.2. Scanning as sampling 3
2.3. Reconstructing the image 3
2.4. Scanning artwork 4
2.5. Combining scanned text with pictures 5
2.6. 2-level electronic images 5
2.7. Multi-level digital images 6

3. The spectral contributions of picture detail 6

3.1. Stationary detail 6
3.2. Moving detail 6

4. The causes of aliasing 7

4.1. 25 Hz flicker (interlace twitter) 7
4.2. Stepping or 'jaggies' 8
4.3. Movement portrayal 10
4.3.1. Pictures from electronic cameras 10
4.3.2. Pictures from film 10
4.3.3. Electronic pictures 10

5. Filtered text images 11

5.1. Source of data 11
5.2. Choice of filter function 12
5.3. Practical investigation 12
5.3.1. Equipment used 12
5.3.2. Display gamma 12
5.3.3. Filters investigated 12
5.3.3.1. Filters with monotonic step-responses 13
5.3.3.2. Undersampling without filtering 13
5.3.3.3. Filters with non-monotonic step-responses 13
5.3.4. Combined filtering 16
5.3.5. Number of bits required 17
5.3.6. 'Pixel-fit' 17
5.3.7. Inter-character spacing 18

6. Use of multi-level filtered text 19

7. Conclusions 19

8. Acknowledgements 20

9. References 20

10. Appendix

Glossary of typographic terms 20

© BBC 2003. All rights reserved. Except as provided below, no part of this document may be reproduced in any material form (including photocopying or storing it in any medium by electronic means) without the prior written permission of BBC Research & Development except in accordance with the provisions of the (UK) Copyright, Designs and Patents Act 1988.

The BBC grants permission to individuals and organisations to make copies of the entire document (including this copyright notice) for their own internal use. No copies of this document may be published, distributed or made available to third parties whether by paper, electronic or other means without the BBC's prior written permission. Where necessary, third parties should be directed to the relevant page on BBC's website at http://www.bbc.co.uk/rd/pubs/ for a copy of this document.

ELECTRONIC GENERATION OF TEXT FOR TELEVISION

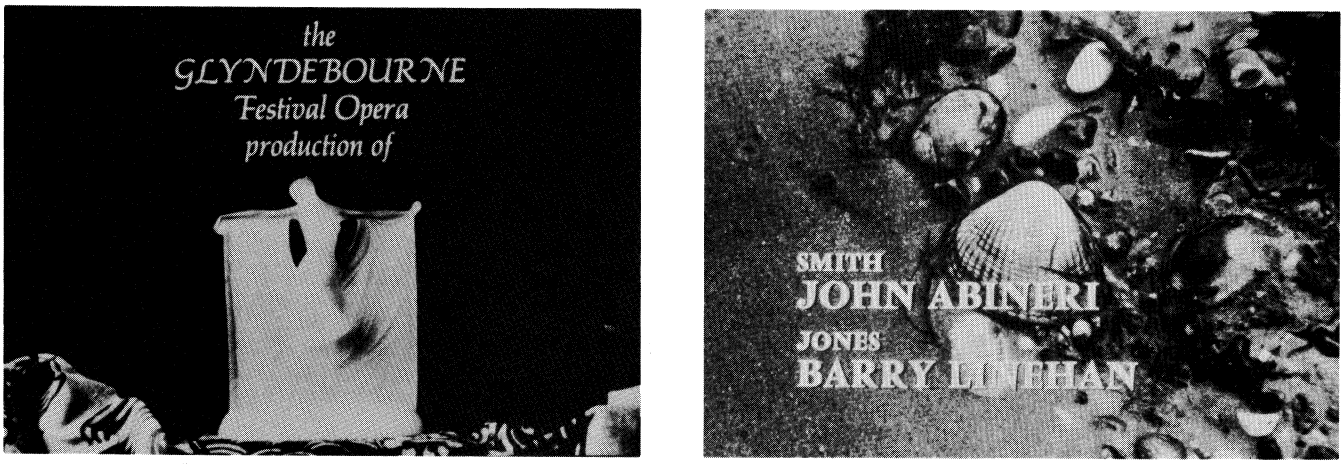
N.E. Tanton, M.A., C.Eng., M.I.E.E., M. Inst.P.

1. Introduction

Most broadcast television programmes include sequences in which the written word appears incorporated into the picture signal. Titles, credits and any explanatory captions within the programme are generated either from artwork or by means of electronic character generators or computer device (Fig. 1).

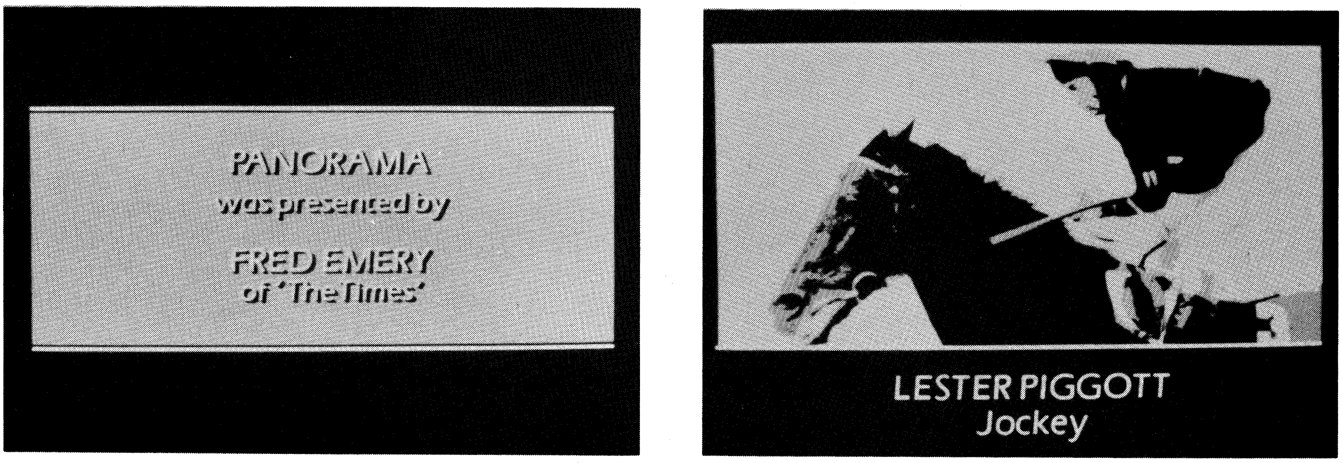
The traditional method of making text in vision involves preparing transparent or opaque artwork and scanning it to form an electronic picture signal. The scanned brightness of the artwork is represented in the video signal by a continuous voltage whose magnitude is a function of the transmission density or reflectivity of the artwork. Transparent artwork is usually in the form of high contrast 35 mm photographic slides which are scanned using conventional slide-scanners. Opaque artwork, which usually comprises printed or phototypeset characters mounted on card, is scanned using a television camera.

Electronic character generators, however, produce an electronic text signal directly without original artwork. Using simple digital techniques they often attempt to represent typographic characters by a two-level pixellated image in which one signal level represents character detail and the other signal level represents background (no character). Individual pages of text are built up character-by-character under control from a keyboard (or more rarely from a remote computer).



Titles derived from conventional artwork

Credits from artwork



Captions from an electronic character generator

Annotated diagrams from a computer graphics device

Fig. 1 – Examples of the use of text in television programmes.

Whether from scanned artwork or character generator, the electronic text signals are subsequently processed and presented as input to circuitry which combines the text with the background picture signal.

In the BBC Television Service, the use of character generators has often been restricted to those areas of programme output (such as Sports and Current Affairs) in which operational considerations of speed of operation or budgetary constraint have dictated their use. In general, television graphic designers maintain strong preferences for the traditional methods of television typography in which real artwork is scanned. The reasons for these preferences are largely aesthetic ones : scanned artwork can produce an image which is considerably more pleasing, in terms both of typographic and of image quality, than any text image produced by current electronic character generators.

Working with photoartwork, the graphic designer has a very wide choice of fount style and character size; this is extremely important if he is to be free to set or to suit the mood of a particular programme. In addition, he has complete freedom over the relative positioning of characters both within the display area and one character with another. Finally, if the text image is properly combined with the background picture signal (see section 2.5) the displayed image maintains the letterforms of the original artwork and is largely free from the unpleasant distortions of flicker on horizontal detail and of stepping on curved and inclined character strokes.

By contrast, the repertoire of fount styles and sizes available on present commercial character generators is very small. There are also restrictions on the absolute and relative positions of characters in the text image, restrictions resulting mainly from the 2-level sampled form of the image and from the implementation of the generator. The letterforms of 2-level character founts (constrained by the relatively low-resolution of broadcast television) are often badly represented and bear only passing resemblance to the eponymous photosetting founts. Finally, when displayed, the text image exhibits the aforementioned distortions of 25 Hz flicker and stepping together with marked movement judder and modulation of character stroke width when captions are rolled or crawled.

There is no reason why electronic character generators should not produce aesthetically acceptable letterforms using digital character representation; the present limitations are largely due to the use of only 2 signal levels for that representation. This Report examines the technical reasons behind the disparity in image quality between scanned artwork and present-day character generators and describes methods of electronic character generation which result in an image quality equal to, or better than, that available using the traditional methods of television typography. The conclusions are equally applicable to other forms of synthetic picture generation including raster-scanned computer graphics images.

2. The effects of scanning

2.1. Sampling theory

When a continuous signal is sampled in the time domain, the spectrum of Fourier components representing the original signal in the frequency domain (Fig. 2(i)) is repeated at harmonics of the sampling frequency f_s (Fig. 2(ii))[1].

The continuous signal may be reconstructed from the samples by filtering the sampled signal with a low-pass filter (Fig. 2(iii)) whose characteristic is such as to isolate the baseband spectrum from its harmonics (Fig. 2(iv)). In principle, an 'ideal' filter characteristic, with the rectangular shape shown in Fig. 2(iii), would pass unattenuated all frequency components up to half the sampling frequency (i.e. $f_{s/2}$) and suppress completely all components beyond this frequency (see below*).

If the continuous signal prior to sampling contains frequency components above half the sampling frequency, (Fig. 2(v)), the repeated spectra overlap in the region around $f_{s/2}$ (Fig. 2(vi)). When reconstructed from the sampled signal using the 'ideal' low-pass filter, the resultant signal includes spurious energy from the high frequency components of the repeated spectra (Fig. 2 (vii)). This spurious energy results in signal distortions known as 'aliasing'. The unwanted overlapping components could be removed by reconstructing the signal with a filter of lower cut-off frequency (Fig. 2 (viii)) at the expense of those wanted high frequency components which were present in the original signal in the region of potential spectral overlap.

---

* In practice, it is desirable to use filters with practicable rates-of-cut and with finite impulse responses, unlike the 'ideal' filter of Fig. 2(iii). In order to leave a margin for real filters, it is usual to sample at just above twice the frequency of the highest component in the continous signal.

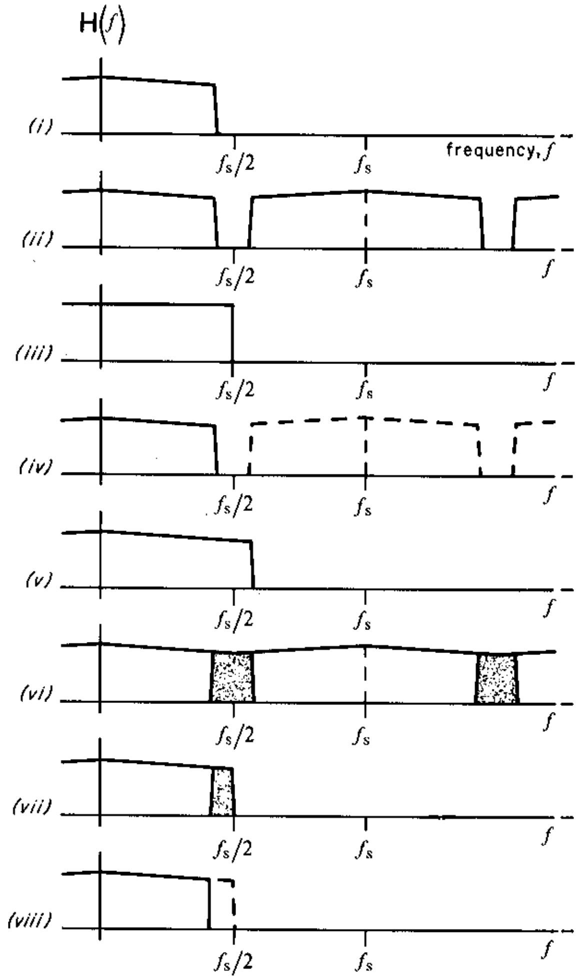


Fig. 2 – One-dimensional spectrum of a continuous signal, its sampled form and of a continuous signal reconstructed from those samples.

(i) Spectrum of original continuous signal

(ii) Spectrum of (i) after sampling at f_s

(iii) Characteristic of the 'ideal' reconstructing low pass filter

(iv) Spectrum of the continuous signal reconstructed with the filter of (iii) from the sampled signal (ii)

(v) Spectrum of continuous signal with components beyond $f_{s/2}$

(vi) Spectrum of (v) sampled at f_s

(vii) Spectrum of continuous waveform reconstructed from (vi) with ideal filter (iii)

(viii) Spectrum of continuous waveform reconstructed with a lpf of lower cut-off frequency

To avoid aliasing, it is necessary to pre-filter the continuous waveform prior to sampling in order to remove those spectral components (at the beyond half the sampling frequency) which would give rise to aliasing distortions.

It can be seen that a discontinuous signal cannot be perfectly represented by a sampled waveform because its spectral components extend ad infinitum. Furthermore, when the sampled waveform is reconstituted as a 'continuous' signal by the low-pass filter, the discontinuous input signal will be represented by a band-limited signal transition with finite rise-time.

2.2. Scanning as sampling

The brightness of an optical image perceived by a television camera tube or telecine photomultiplier is a three-dimensional function of horizontal position, vertical position and time. This brightness function has a corresponding three-dimensional spectrum describing the energy distribution of spectral contributions due to the picture detail, in terms of horizontal and vertical spatial frequencies and of temporal frequencies resulting from motion within the scene. The extent and shape of this spectrum is determined by the scene detail and the optics; it will be limited by practical factors such as aperture loss (see below*) and camera integration.[2]

When the analysing raster pattern of the camera or flying-spot tube scans this optical image, the three-dimensional brightness function is converted to a one-dimensional signal of voltage varying with time in which the brightness is defined only at discrete vertical positions and time instants determined by the scan. The image has, therefore, been sampled in two dimensions. As a result, the three-dimensional spectrum of the 'optical' image is repeated in the frequency domain at integral multiples of the vertical and temporal scanning frequencies. In a 625-line, 50 fields/second scanning system, the spectrum is repeated at multiples of 625 cycles/picture height and 50 Hz (Fig. 3(ii)). In addition, the 2:1 interlace structure also causes the spectrum to be repeated at odd multiples of (25 Hz, 312.5 c/ph) Fig. 3(iii)). This structure of repeated spectra also obtains when a synthetic image is represented in scanned form, as with electronic character generator and other raster-scanned electronic graphic images.

---

* Throughout this Report, vertical spatial frequency will be expressed in cycles per full picture height (i.e. including picture blanking). Thus the Nyquist limit of vertical spatial frequencies is expressed as 312.5 c/ph for a 625-line raster.

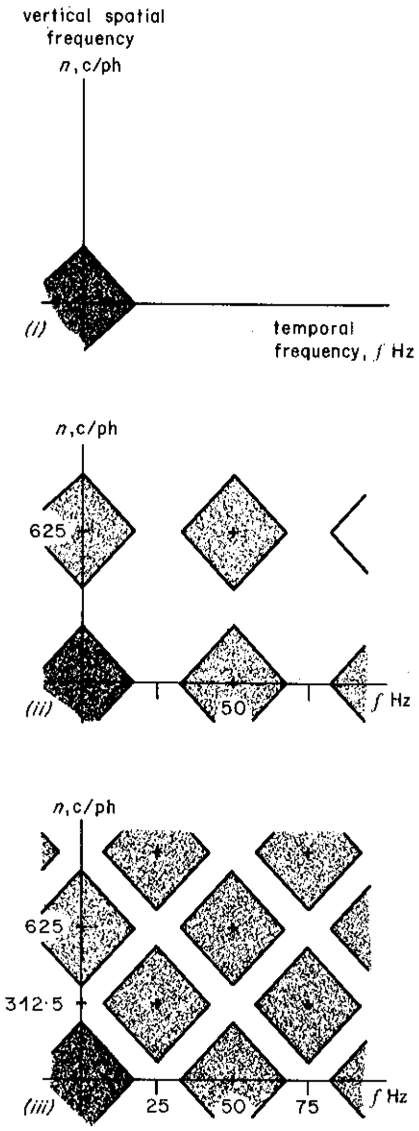


Fig. 3 — Two-dimensional spectrum of an optical image and its scanned form.

(i) Spectum of optical image (as section 3-D spectrum).

(ii) Spectrum of image scanned with 625-line 50 fields/second sequentially scanned raster.

(iii) Spectrum of image scanned with 625-line 50 fields/second 2:1 interlace raster.

2.3. Reconstructing the image

The image sampled by the scanning process is reconstructed by the television display as a continuous three-dimensional (x, y, t) brightness signal. The display tube filters the spatial and temporal components of the sampled image with a low-pass characteristic because the display spot has a finite size and the display phosphor a lag-like temporal response. This low-pass filtering is somewhat imperfect; there remain some direct harmonic components in the reconstructed picture of which the most noticeable is usually the field-rate flicker visible in plain areas of high brightness.

If there is spectral overlap in the spectrum of the scanned image, aliassed components remaining in the filtered signal will appear as distortions on the displayed picture. It is the presence of these alias components in picture signals which give rise to 25 Hz flicker (twitter), stepping and movement portrayal problems such as judder. Spectral overlap (and thus aliasing distortions) can be reduced or removed by suitably prefiltering the image prior to forming a scanned version of it. The disparity in text image quality between scanned artwork and most present-day electronic character generator images exists because the artwork image is effectively prefiltered by the scanning device.

2.4. Scanning artwork

The spatial definition of scanned artwork is very high, limited by photolithographic film grain size for example. Nevertheless, the displayed text image which represents it is considerably less sharp because the scanning, transmission and display processes impose physical limits on the picture detail which may be displayed.

There is inherent loss of detail in the scanning process because the signal contributions of light from each point on the artwork are spread over a larger finite area (or 'aperture') of the electronic image. A simple example of such loss occurs when the camera lens is defocussed.

This loss of detail caused by spreading is called 'aperture loss' and results in a softening or blurring of high-frequency picture detail (including abrupt transitions). The loss is attributable to several physical causes both in the scanning and in the display processes; these causes include:

(i) lens imperfections (flare, low resolution etc.)

(ii) image spread in the camera-tube target or film emulsion;

(iii) scanning spot-size in the camera tube or slide-scanner flying-spot tube;

(iv) scanning spot-size in the display.

Partial correction for aperture loss in the scanning process is usually applied by selectively boosting the high frequency components of the picture signal; these components represent fine picture detail. Such correction is limited in scope, partly because there are still fundamental limits to the detail which may be transmitted and displayed by the television system and partly because this detail may be restored only at the expense of increase electrical noise^2.

These limits to the reproducible detail mean that the abrupt changes of reflectance or transmission density which mark the boundaries of typographic characters in the artwork cannot be reproduced as perfectly abrupt voltage transitions of the video signal or as perfectly abrupt changes of brightness in the displayed picture. Instead, these changes are represented by more gradual, band-limited transitions. Fig. 4 shows the video signal from the slide-scanner in which an abrupt density change in a transparency has been scanned.

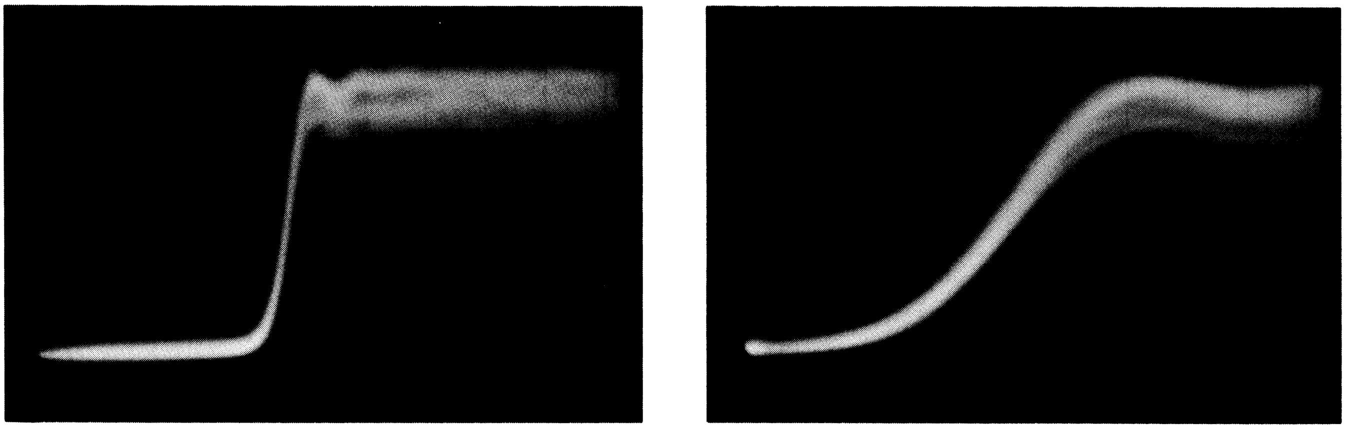


Fig. 4 – The step-response of a television slide-scanner. (both traces showing signal after after-glow correction in linear signal domain).

Small changes to the position of detail within the scanned artwork result in changes to the timing of voltage transitions in the scanned signal and thus to the position of the displayed image detail. In this way the infinite control of the absolute and relative positions of characters which is possible within titling artwork is matched by a continuously variable adjustment of position in the displayed picture.

Aperture loss in scanning is equivalent to prefiltering the image prior to sampling it by scanning. The loss is advantageous in that spectal overlap in the scanned signal is significantly reduced.

2.5. Combining scanned text with pictures

Text signals derived directly from scanned artwork are usually free from most aliasing distortions when displayed. In some studio vision mixers, however, the text signal from camera or slide-scanner is sliced to form a keying signal with which to switch colour into the background video signal.

When sliced, the band-limited waveform is converted to an un-band-limited 2-level signal with abrupt transitions; the advantages of prefiltering are immediately lost. The resulting characters lose much of their fine detail (especially around serifs) and display all the aliasing distortions typical of 2-level graphics together with an unpleasant vacillation of character edge position both on horizontal and verticals. This caused by electrical noise which perturbs the slicing decision level.

The keying process, in essence, involves multiplying the background video signal by the 'key'. In the frequency domain the spectrum of the background video signal is convolved with that of the 'key'. If the keying signal has unlimited spectral components (as is the case in a 2-level signal) then the spectrum of the combined image signal will also have unlimited components and aliasing distortions result.

If, however, the keying signal is filtered in the same fashion as the video signal, the combined image signal will also be band-limited. By combining scanned text and background video in a circuit such as shown in Fig. 5 using a linear multiplication process, the prefiltering advantages

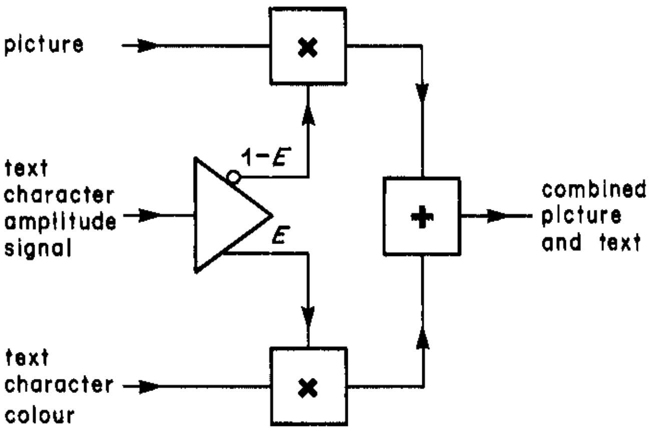


Fig. 5 – Combining pictures with text.

provided by the scanning devices are maintained in the output image.

The principles of soft keying are applicable equally to analogue scanned text, synthetic images or to the multi-level digital character generator techniques described below; it is indeed essential to use 'soft keying' to preserve the image quality of prefiltered text signals.

2.6. 2-level electronic images

Two-level pixellated images, such as those formed by most current character generators, inevitably suffer from aliasing distortions due to the spectral overlap described above. These distortions will be evident even if 2-level text and video are combined in a 'soft-keyer'. Additionally, because the image is in a horizontally sampled form there is a quantisation of edge position which places limits on the accuracy of letterform of each character and on the control of inter-character and inter-line spacing which may be achieved. Some commercial character generators have made use of very high effective horizontal sampling frequencies to improve the horizontal resolution, sometimes improving the letterform without changing the spacing accuracy. As the vertical resolution is prescribed by the number of scan-lines this approach is not practicable in the vertical direction (see Section 5.3).

2.7. Multilevel digital images

When an abrupt transition is low-pass filtered, the resulting response is represented by a continuum of signal levels intermediate between the two levels representing the input transition. In order to describe a filtered image in digital form, it is clearly necessary, therefore, to represent the image detail with digital words of several bits or more.

By forming a suitably filtered multi-level image it is possible to generate text and other synthetic pictures digitally with a displayed image quality comparable to, or better than, that of scanned artwork. (Because such filtering reduces aliasing distortions it is sometimes known as 'anti-aliasing'). The final displayed image quality depends on the level of filtering and represents a balance between residual aliasing distortions and those picture impairments which are caused by the filtering itself. Such impairments include loss of detail or excess signal overshoot resulting in fringeing around edge detail.

As the nature of the filtering is important, it is appropriate to identify the spectral contributions made by various forms of picture detail and to relate particular aliasing distortions to the spectral components which cause them.

3. The spectral contributions of picture detail

3.1. Stationary detail

Stationary image detail results in spatial frequency contributions with no temporal component – picture detail orientated vertically as components along the horizontal frequency axis and detail orientated horizonally as components along the vertical frequency axis. Diagonal detail gives rise to spatial frequencies which may be resolved into horizontal and vertical frequency components (Fig. 6).

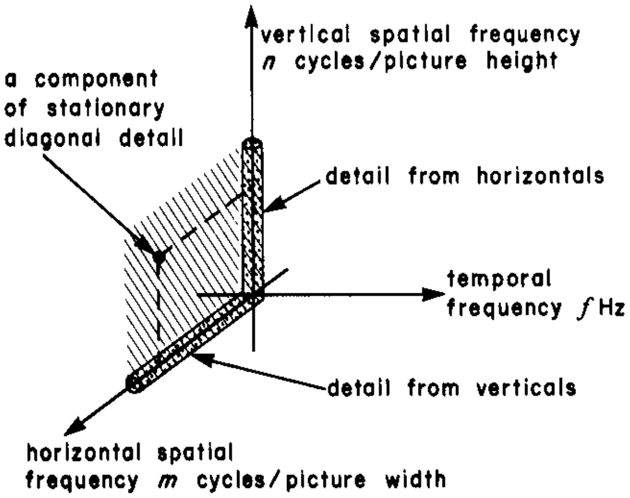

vertical spatial frequency
n cycles/picture height
a component of stationary diagonal detail
detail from horizontals
temporal frequency f Hz
detail from verticals
horizontal spatial frequency m cycles/picture width

Fig. 6 – Spectrum of stationary picture detail.

3.2. Moving detail

If the image detail moves, the spatial

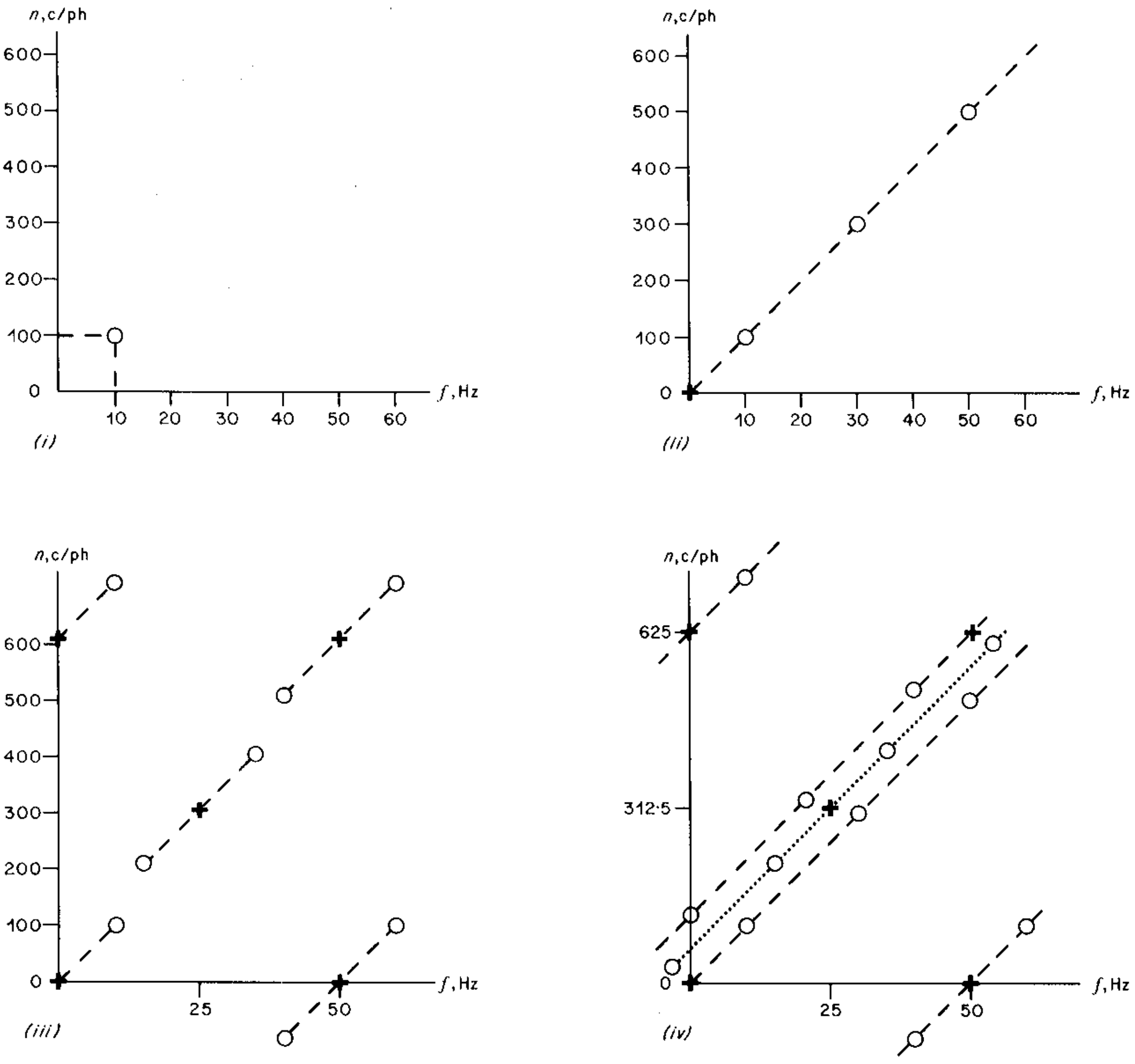


Fig. 7 — Spectrum of moving picture detail.

(i) Vertical sinusoidal grille (100 c/ph moving at 0.1 ph/sec). (ii) Vertical square-wave grille.
(iii) Vertical sinusoidal grille as scanned. (iv) Vertical square-wave grille as scanned.

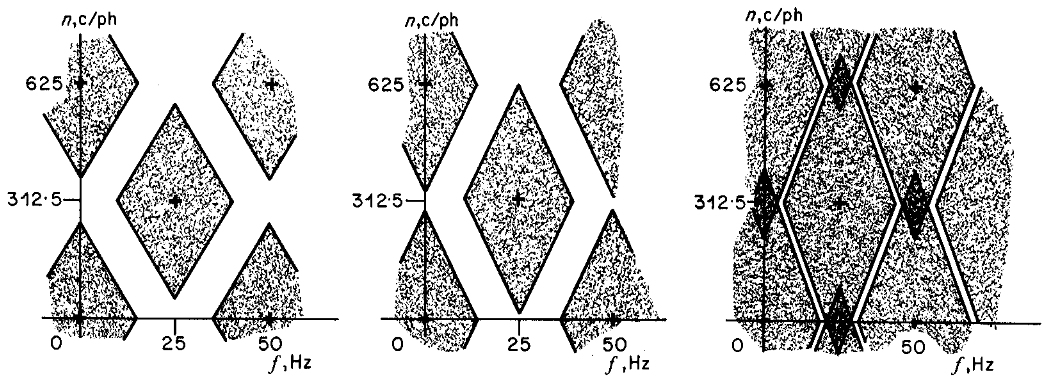
frequency contributions due to the detail acquire temporal components whose temporal frequencies are proportional both to the magnitude of the movement and to the spatial frequency of each spatial component. Fig. 7(i) shows the baseband 2-dimensional spectrum of a sinuisoidal grille version of spatial frequency 100 cycles/picture height moving vertically at a speed of 0.1 picture heights per second, a speed typical of roller captions. The corresponding temporal component is 10 Hz. As has been mentioned, abrupt steps have spatial components extending ad infinitum; when such a transition moves, an infinite number of spatiotemporal components will result. Fig. 7(ii) therefore, shows part of the spectrum of a square wave grille of the same pitch and motion as for Fig. 7(i). This detail is more typical of electronically generated rolling captions.

When the baseband spectrum is repeated by the scanning process (see above) all the spatiotemporal components are repeated about the repeat sites; Figs. 7(iii) and 7(iv) show some of the repeated components due to the moving grilles of Figs. 7(i) and 7(ii).

4. The causes of aliasing

The three forms of aliasing distortion particularly associated with synthetic pictures are:

(1) 25 Hz flicker on horizontal or near horizontal



(i) (ii) (iii)

Fig. 8 — Spectrum of moving picture detail when scanned.

(i) Moderate vertical detail. (ii) Increased vertical detail. (iii) Excessive vertical detail at and beyond 312.5 c/ph.

detail (interlace twitter);

(2) stepping or 'jaggies' on near horizontal and near vertical edges and on curved strokes;

(3) movement judder and stroke-width modulation on moving images (e.g. roller captions).

4.1. 25 Hz flicker (interlace twitter)

If there is little high frequency vertical detail in the scanned image, the repeated spectra are well separated in the vertical and temporal frequency plane (Fig. 8(i)). As the high frequency vertical detail is increased (e.g. by focussing the camera) the spectra are extended further in the vertical frequency direction (Fig. 8(ii)) until they overlap when there are vertical frequency components at and beyond half the vertical sampling frequency (i.e. beyond 312.5 c/ph). In particular, a purely spatial component at (0, 312.5) is repeated by the interlace spectrum centred on (25, 312.5) to give rise to a purely temporal component at (25,0). In reconstructing the brightness function from the sampled waveform, the display will interpret components from this region of spectral overlap as 25 Hz flicker.

When artwork is scanned by camera or slide-scanner, the aperture loss is usually such that there is relatively little spectral energy at or above 312.5 cycles/picture height; 25 Hz flicker is then not very noticeable. The high frequency spatial components of 2-level synthetic pictures often include significant energy at 312.5 cycles/picture height resulting in visible 25 Hz flicker. It should be noted here that if a synthetic image contains some picture detail comprising an even number of television picture lines of the same brightness than that detail contributes no energy at (0, 312.5) and therefore generates no 25 Hz flicker. For this reason founts are often readjusted or 'redesigned' so that the horizontals (including serifs if appropriate) occupy an even number of picture lines. As a consequence, the 'redesigned' fount often differs visibly from the phototypesetting letter forms on which it is based. This adjustment is totally unnecessary with properly filtered characters.

4.2. Stepping or 'Jaggies'

When the detail is sufficient to cause the spectra to overlap in the vertical frequency direction (Fig. 8(iii)) vertical stepping becomes visible on near-horizontal lines and curved strokes. This stepping (also known as 'the jaggies') (see Fig. 9) results from coherent energy components in the baseband spectrum beyond half the vertical sampling frequency (312.5 c/ph) being folded-back or aliassed into the baseband region by the spectral repetition. If the amount of overlap is significant then this aliassed energy is well spread over the baseband spectrum and cannot be separated out when the image is reconstructed by the display.

If the image is also sampled in the horizontal direction (as with the pixellated images of electronic character generators and most electronic or computer graphics devices) the baseband spectrum is repeated not only in the vertical spatial and temporal frequency plane but also along the horizontal spatial frequency axis. In this case the spectra are also centred at integral multiples of the horizontal sampling frequency.

The presence of horizontal spatial frequency components above half this horizontal sampling frequency will give rise to horizontal aliasing distortions. The sampled horizontal detail is reconstructed as a continuous electrical waveform by a horizontal low-pass filter before the whole image is reconstructed by the display. Horizontal aliasing is manifest as horizontal stepping on near-vertical lines and curved strokes.

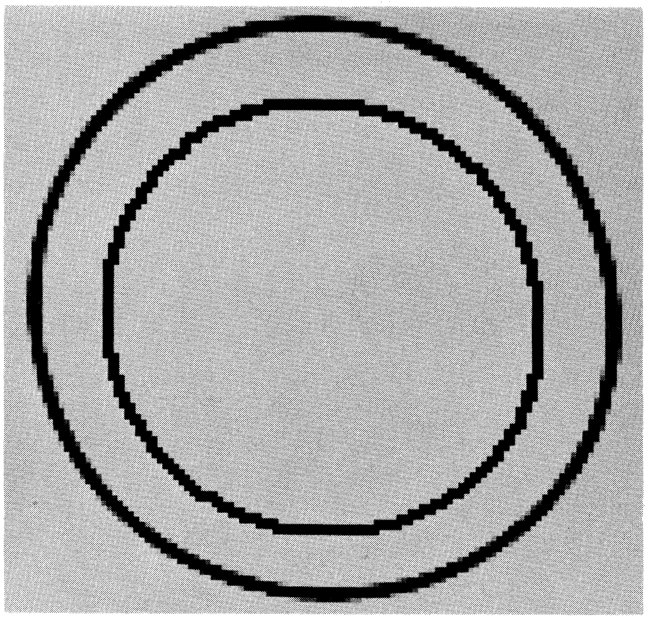


Fig. 9 – Example of stepping or 'jaggies'; inner circle without antialiassing, outer circle with antialiassing (expanded x 4).

Fig. 10(i) shows a typical 2-level typographic character as photographic artwork. Fig. 10(ii) demonstrates the letterform when the artwork of 10(i) has been coarsely sampled vertically and horizontally with equal resolution. The stepping is the direct consequence of undersampling the image of 10(i) without filtering prior to sampling. As mentioned in Section 2.6. some electronic character generators make use of high horizontal sampling frequencies in order to reduce the apparent horizontal step size (Fig. 10(iii)). There is no opportunity for doing the same vertically with 2-level images because the vertical resolution is prescribed by the raster. The stepping does become less visible as sample rates increase (Fig. 10(iv)) but the character sizes representative of practical television typography (i.e. between about 16 and 60 pictures lines height for a capital M) do not allow much margin for masking the steps in this way.

It should be noted that the stepping is a direct consequence of the 2-level representation of the image irrespective of pixel dimensions.

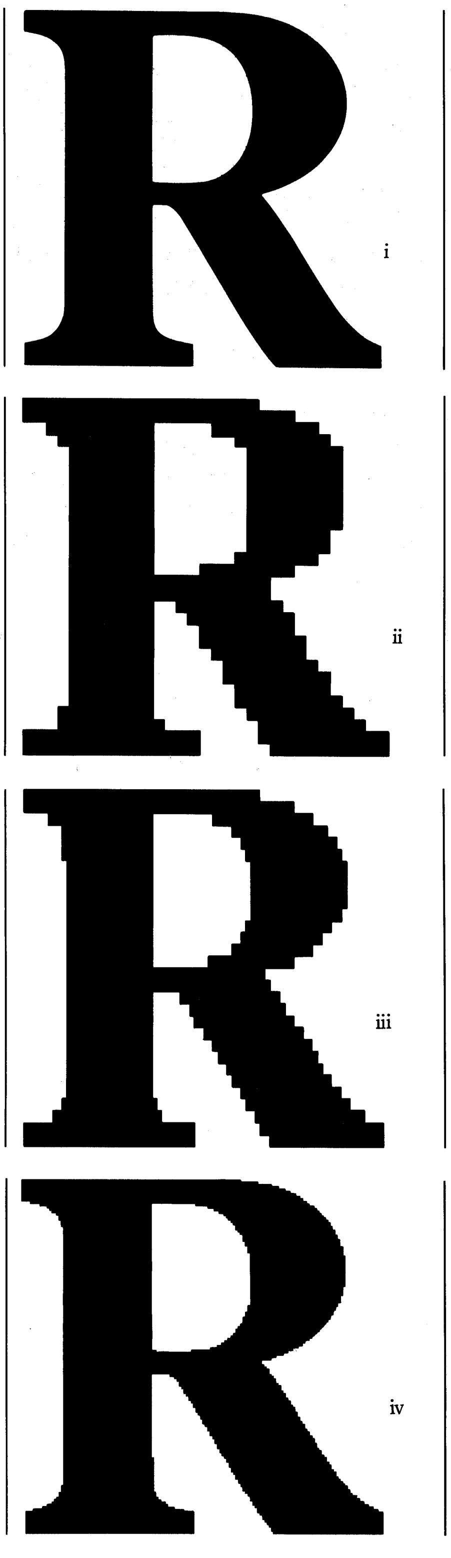


Fig. 10 – Typographic characters as represented in two-level form.

(i) High-definition photolithographic quality.
(ii) Coarsely quantised on 30 x 30 matric of pixels.
(iii) Quantised with same vertical resolution as (ii) but with horizontal resolution x 2.5.
(iv) Quantised on 150 x 150 matrix.

Scanned artwork on the other hand has very little energy beyond 312.5 c/ph and there is therefore very little vertical aliasing. Unless the scanned signal is then inopportunely sampled horizontally (e.g. for some digital processing such as picture manipulation) there will be no horizontal aliasing. Only when scanned artwork is sliced (e.g. to form a 2-level keying signal) will stepping be noticeable (see Section 2.5).

4.3. Movement portrayal

The detail in the unscanned image of a stationary scene may be described by a series of Fourier components of spatial frequency. As has been noted (Section 3.2), when the scene moves each spatial frequency will acquire a temporal component; the greater the spatial detail or the movement speed, the greater will be the resultant temporal components. With a sharply-focussed scene these temporal components may be quite large even for modest movement.

Fig. 11(i) shows the two-dimensional spectrum of an arbitrary stationary object; in the absence of motion there are no temporal components. Fig. 11(ii) shows the spectrum when the object moves with a vertical component of velocity. The spectral repeat caused by scanning gives rise to a spectrum shown in Fig. 11(iii). Due to the overlap the display is unable to distinguish between the baseband components of fast-moving low frequency detail and the interlace spectral components of slow-moving, high-frequency detail. This results in temporal aliasing which is manifest as uneven movement portrayal (e.g. judder), a modulation of the position of moving edge detail (on rolling captions for example) and the well-known effect of wagon-wheels appearing to rotate backwards.

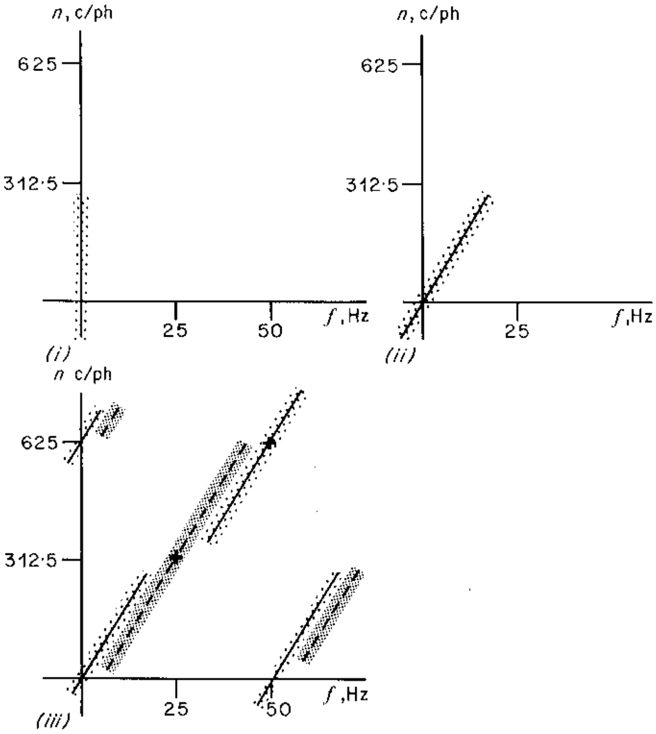


Fig. 11 – Two-dimensional spectrum of moving detail.

(i) Stationary object.

(ii) Moving with vertical component of motion.

(iii) Scanned version of (ii).

4.3.1. Pictures from electronic cameras

Although temporal aliasing is a fact-of-life with broadcast television the effects are mitigated by the use of electronic television cameras. Aperture loss and the field integration of a camera tube serve to attenuate the higher spatial and, to a lesser extent, temporal frequencies when the optical image is scanned. Fig. 12 shows the approximate spatial and temporal characteristic of an electronic camera*. The attenuation afforded to high spatial frequencies by aperture loss probably has a greater effect in reducing temporal aliasing than the temporal attenuation caused by field integration in the target of the tube.

4.3.2. Pictures from film

Film cameras on the other hand provide sharp pictures and sample the scene only every 24 or 25 times a second. With less spatial frequency loss and integration over only half the frame period (180° shutter angle) temporal aliasing is more severe, as exemplified by the wagon-wheels effect and judder on rapid camera movements.

4.3.3. Electronic pictures

Electronically generated pictures are often totally unfiltered spatially or temporally. The high spatial frequency components present thus give rise to high temporal components when the image is made to move, resulting in severe temporal aliasing. The movement judder and stroke-width modulation evident on digitally generated roller captions provide obvious examples of this aliasing, effects which are considerably less noticeable on scanned artwork captions.

---

* based on a model by Dr. J.O. Drewery.

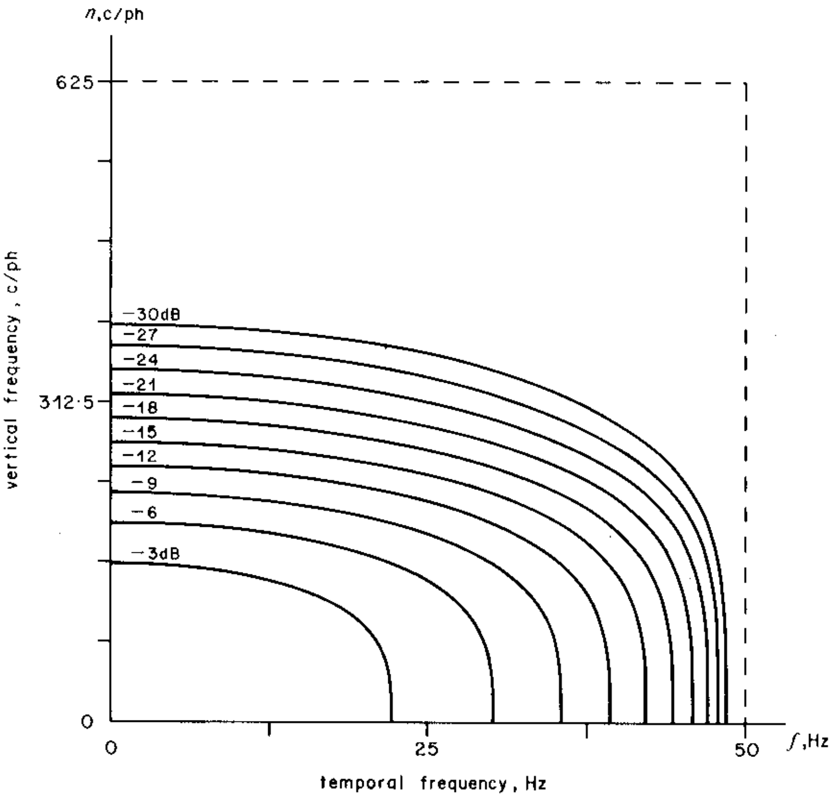


Fig. 12 — Approximate vertical spatial and temporal characteristic of an electronic camera.

5. Filtered text images

In order to minimise aliasing, electronic images should be generated in filtered form. From the preceding discussion it is clear that spatial frequency components at or beyond half the vertical and horizontal sampling frequencies should be attenuated to mitigate interlace twitter, stepping and, to a lesser extent, movement distortions. If the image is to be seen to move then some temporal filtering may be desirable to further reduce temporal aliasing effects. One possible two-dimensional spatio-temporal filter template suitable amongst other purposes for electronic graphics is shown in Fig. 13.

Although no work has yet been done on spatio-temporal filtering for rolling or crawling captions, the spatial filtering of typographic images has been investigated to generate digital text with an image quality comparable to that of scanned artwork.

5.1. Source of data

Master fount data is available in digital form from a large number of typefoundries. The data is derived from reference artwork, carefully prepared by typographers, which is digitised to two levels at very high resolution. The data is edited to make minor modifications and to eliminate obvious digitising errors and then it is stored character-by-character on a suitable magnetic storage medium. The exact format of stored data varies from source to source (e.g. run-length, vector, feature description etc.

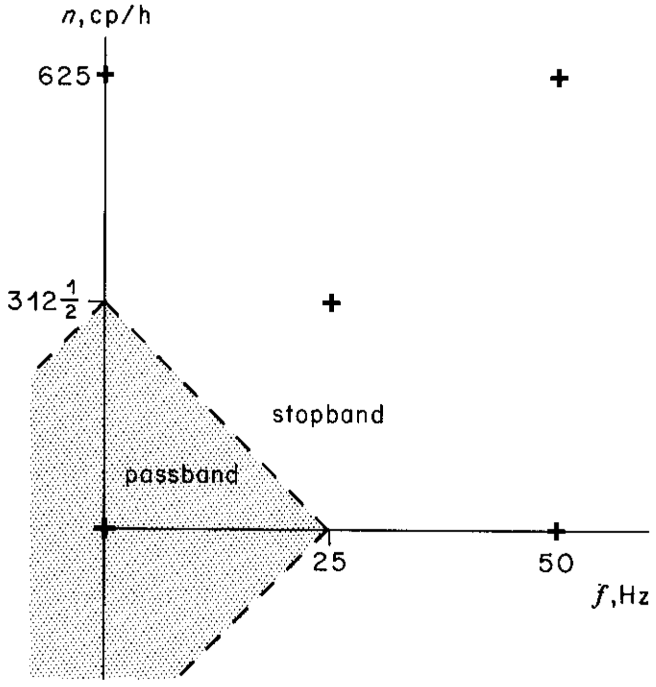


Fig. 13 — A possible two-dimensional spatio-temporal filter characteristic for electronic graphics.

As supplied, this master data is unsuitable for direct use in television because it is in two-level form, it may not be in scan line format and it is of far higher resolution (typically greater than the equivalent of 1000 lines character height for a capital M) than is required for normal television typography, where character heights of 16 – 60 lines are more appropriate. It is thus essential to undersample the character images and if necessary to convert them to scan-line format for display.

Using software, the master data is low-pass filtered in the two spatial directions and the resultant multi-level signal resampled at the requisite output resolution. The filter characteristics are scaled to suit the ratio between the input and output sampling resolutions and are chosen according to the criteria described above. In this way well-shaped characters may be generated from the master data for any character size practicable for broadcast television*. A particular advantage of multi-level text prepared in this way is that all features of the master data make a contribution to the final image – within the limitations of the television display system, the displayed character should therefore display the correct proportions of strokes, serifs and counters as are embodied in the master data.

5.2. Choice of filter function

The complexity of low-pass filter used to process the master data is less important than the image quality because the filtering is performed by software and need only be performed once per character for a given fount style and output character size. In balancing the factors influencing image quality the following points apply:

(i) too soft a filter will give indistinct characters with degraded edges and serifs, filled-in counters and soft 'nicks' (see Appendix).

(ii) insufficient filtering will result in aliasing distortions.

(iii) too sharp a low-pass filter will result in fringeing around and within character strokes, due to the unbounded step response. (Although signal overshoot will give rise to a distinct image, the fringeing caused by sharp-cut low-pass filters may be typographically unsatisfactory).

5.3. Practical investigation

5.3.1. Equipment used

Master fount data provided by The Monotype Corporation has been filtered and displayed using the electronic graphics system known as 'Flair'. This equipment comprises a microcomputer-controlled pixel-addressable frame store in which a digital image is created and stored and from which it can be displayed. With each pixel is associated 8-bits of storage, the displayed image being 702 pixels wide by 575 pixels (lines) high. Monochrome images may thus be displayed with 256 grey levels and the horizontal pixel resolution is 74 ns.

Purpose-written software running on the microcomputer has enabled the effects of different filtering functions on image quality and letterform to be investigated. The program written in PASCAL allows vertical and horizontal filter functions to be independently specified in analytic terms by defining either the frequency characteristic or the impulse response of the filter function. The one-dimensional frequency characteristic and the impulse response of each filter is plotted on the display for reference and a filtering aperture calculated. After specifying the master fount-data file and the character to be taken from that file the filtering apertures are convolved with the master character data and the resulting multi-level letterform displayed.

5.3.2. Display gamma

The output from the frame-store of 'Flair' is fed to digital-to-analogue converters via a look-up table which in this particular use enables the picture signal to be gamma-corrected to suit the display. Incorrect gamma-correction can greatly reduce the effect of filtering – in the limit a high gamma would effectively result in a 2-level text image and the filtering would be void. On the other hand, if the text signal is being used in the circuit of Fig. 5 as a soft-key control then it should be totally linear (i.e. gamma = 1).

5.3.3. Filters investigated

The filtering functions investigated may be simply divided into two categories, those with monotonic and those with non-monotonic step-responses. Filters with monotonic step-responses do not give rise to pre- and post-transition signal overshoot – there is no ringing on a step. Although inherently less sharp-cut in their frequency characteristic, these filters are of particular interest for two reasons.

---

* The author acknowledges the likelihood that at smaller character sizes the television graphic designer will probably use a different 'weight' of the fount.

First there is no fringeing around character detail due to the lack of overshoot on edges – this may make letterforms more acceptable typographically (see also below). Second, present practical soft-key circuits (such as shown schematically in Fig. 5) usually operate on analogue input signals for which a 'negative' background signal has no meaning. For such circuits monotonic step response filter functions are appropriate for processing the key signal whereas the sharper cut-off filter functions with non-monotonic step responses may be practicable in future digital keying circuits designed to accomodate signed keying signals.

5.3.3.1. Filters with monotonic step-responses

(1) Area averaging

Area-averaging involves summing equal contributions from all the input data samples within a prescribed area centred on the output sample site. The appropriate impulse response and its frequency characteristic is shown in one dimension in Fig. 14 (i,ii). The resultant letterform is rather soft but flicker may be eliminated by choosing the height of the averaging aperture to encompass exactly two output picture lines – Fig. 14(iii) – this causes a zero in the vertical frequency characteristic at 312.5 c/ph. Stepping is just visible due to the incomplete attenuation beyond the first zero in the characteristic.

A character image is shown expanded in Fig. 14(vi). Each pixel in the image has been repeated a number of times horizontally and vertically in order to demonstrate the use of intermediate grey levels.

(2) Linear interpolation

The contribution made by an input data sample to the output sample value can be weighted linearly from being zero at a prescribed distance to a maximum value at the output sample site. The impulse response and filter characteristic of such a linear-interpolation filter is shown in Fig. 15. As with area-averaging, judicious choice of vertical aperture width will result in a zero at 312.5 c/ph and no flicker. The corresponding frequency characteristic will slope more quickly and have more attenuation beyond the first zero than for area-averaging; the resulting image is thus softer but with less stepping.

(3) Filters with an impulse response having a 'raised cosine' transition

If the filter function has an impulse response of raised-cosine shape as shown in Fig. 16, the image quality can be made to be quite sharp but only at the expense of flicker and stepping.

(4) Gaussian, impulse response filters

A filter function with an impulse response of the form $h(t) \alpha \exp(-kt^2)$ has a Gaussian impulse response and Gaussian frequency characteristic (Fig. 17). The scanning spot of a telecine flying-spot or television camera tube has a radial spread-function of Gaussian shape. A Gaussian filter may thus be used approximately to model (in a variables separable form) the non-aperture-corrected image of scanned artwork. If the argument of the Gaussian is chosen approximately to model the vertical frequency characteristic of a television camera tube (i.e. approximately –4 dB at 156.25 c/ph), the filtered character is slightly soft and largely free from aliasing distortion. Decreasing the width of the impulse response results in a sharper image at the expense of flicker and a small amount of stepping (this is true of the Gaussian depicted in Fig. 17).

A Gaussian filter function may represent a practical balance between image sharpness and aliasing for monotonic filter functions.

5.3.3.2. Undersampling without filtering

The generation of 2-level characters by subsampling without filtering can be represented by convolution with a function whose impulse response is unity at the centre of the output sample site and zero elsewhere (Fig. 18).

This was implemented in order to test the software and to allow side-by-side comparisons of 2-level characters with filtered ones. The resultant image predictably suffers from stepping and flicker.

5.3.3.3. Filters with non-monotonic step-responses

Filters with non-monotonic step-responses have pre- and post-transition signal overshoots, producing a sharper image at the expense of fringeing around and within the letterform.

(1) Aperture corrected camera characteristic

The filtering effect of an aperture-corrected

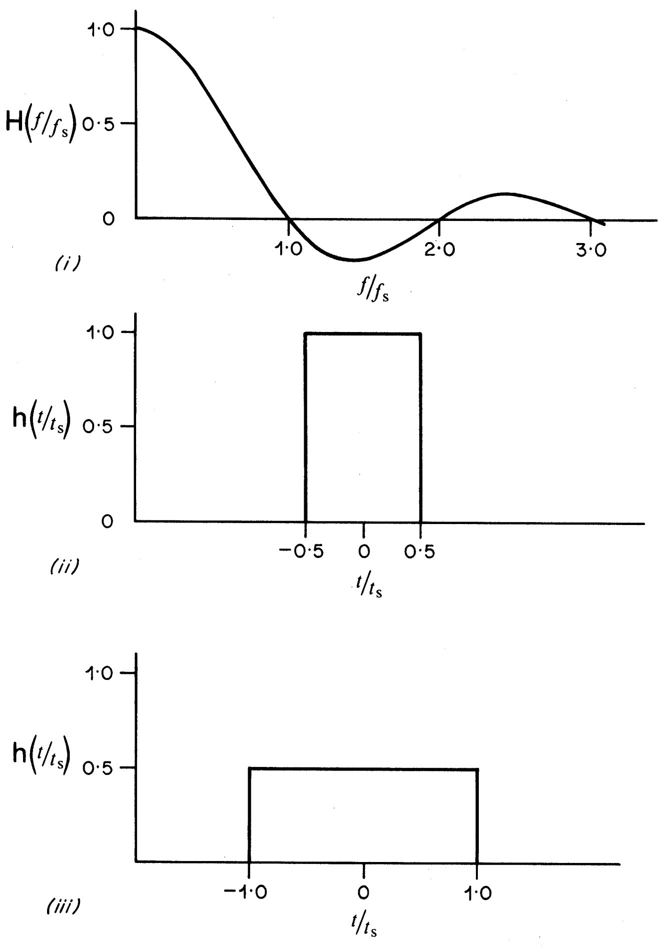
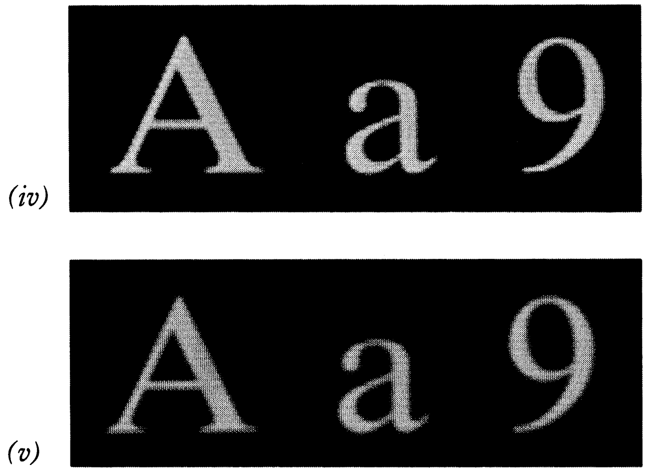
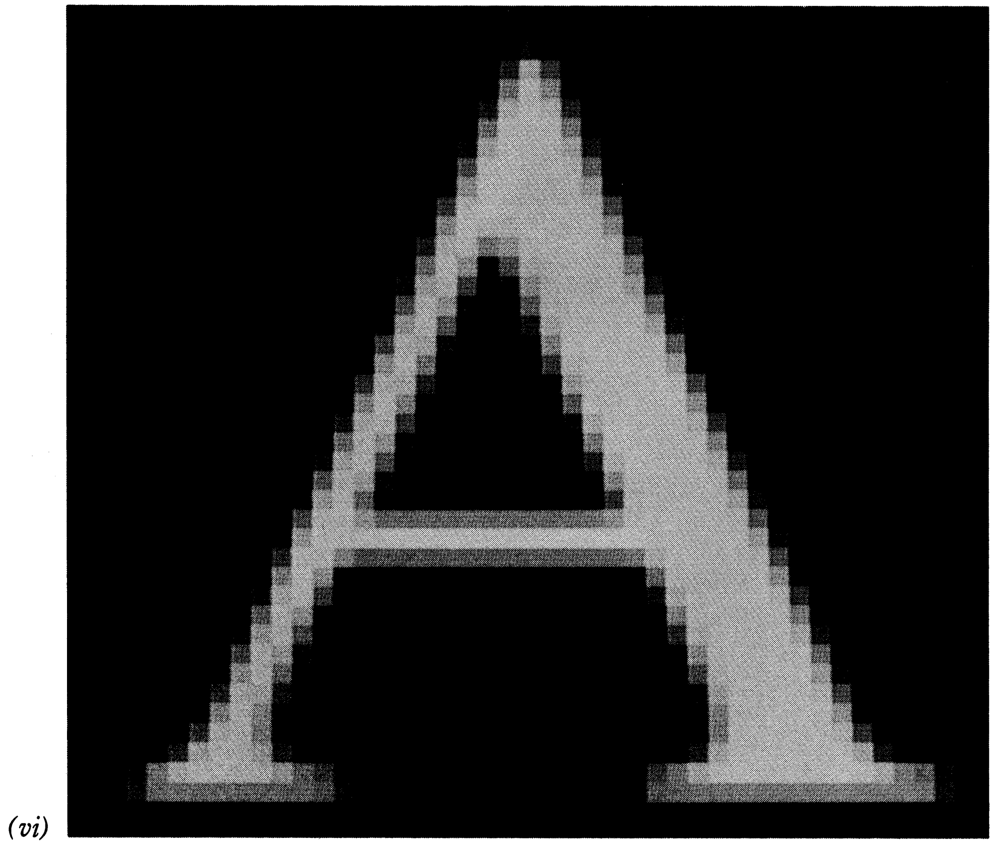


Fig. 14 – Area-averaging filter.

(i) Frequency characteristic of filter a
(ii) Impulse response of filter a (spatial aperture width one sample or line)
(iii) Impulse response of filter b (spatial aperture twice width of filter a)
(iv) Characters filtered with filter a
(v) Characters filtered with filter b
(vi) Capital A filtered with filter a and expanded to demonstrate the use of intermediate grey levels to represent the signal transitions





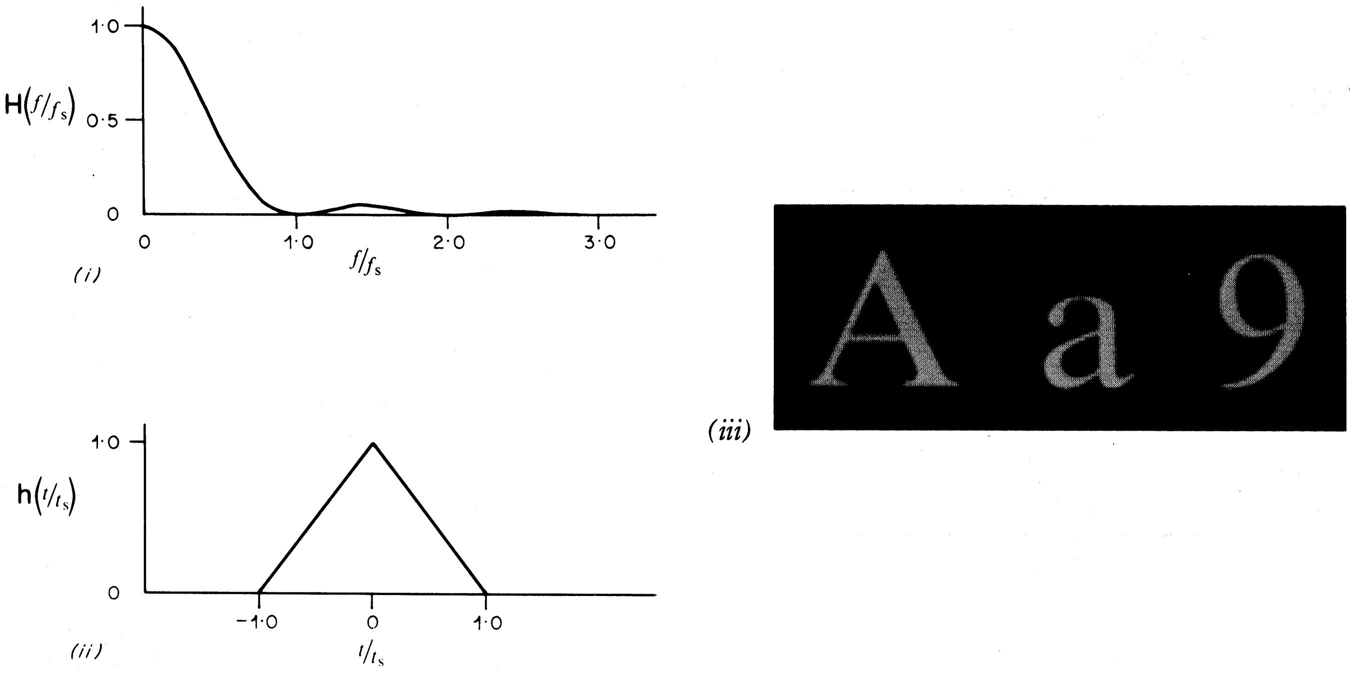


Fig. 15 – Linear interpolation filter.

(i) Frequency characteristic of filter. (ii) Impulse response of filter. (iii) Characters filtered using this filter.

camera scanning real artwork can be synthesised approximately by using a filter function which combines the Gaussian filtering of the scanning spot (see Section 2.4) with the first-order aperture correction usually applied to the camera tube signal. This correction[2] is usually arranged to provide a maximally-flat characteristic fully correcting the aperture loss at 156.25 c/ph vertical frequency. Fig. 19 shows the impulse response and frequency characteristic; the dotted line represents the characteristic of the uncorrected Gaussian filter. The resulting text image is quite sharp, aliasing distortions are hardly visible and the fringeing is restricted in extent to the immediate vicinity of the character edges.

Although the camera-like image is not a perfect simulation as it uses 'variables separable' filtering for the purpose, this filtering provides a useful comparison with which to judge other filtering functions.

(2) 'Ideal' low-pass filter

The 'Ideal' low-pass filter with its sharp-cut characteristic (described in section 2.1 and Fig. 2(iii)) may successfully be implemented in software. Predictably the resulting images are sharp and free from aliasing when stationary but there is very noticeable fringeing (known as 'ringing') around and within the character strokes (Fig. 20).

The fringeing may be reduced in amplitude (at the expense of image sharpness) by choosing filter functions with less sharp-cut characteristics or it may be reduced in amplitude and spatial extent by multiplying the impulse response of the 'ideal' sharp-cut filter with a monotonic function which suitably attenuates all but the fringe closest to detail transitions. This process if called 'windowing' and necessarily produces a less sharp-cut filter characteristic.

(3) Less sharp-cut low-pass filters

One suitable low-pass filter based on the 'ideal' characteristic has a transition band described by one half cycle of a raised-cosine waveform (Fig. 21). As the width of the transition region widens so the fringeing amplitude is reduced but both the amplitude and the spatial extent of the ringing are unacceptable until the image has become quite unsharp.

(4) Windowed impulse response filters

In these filter functions, the impulse response of a fairly sharp-cut filter is further modified by multiplying the impulse response by a weighting function which progressively attenuates the ripples in the impulse response. This modification of the impulse response, therefore, affects the final filter characteristic which becomes the convolution of the original characteristic

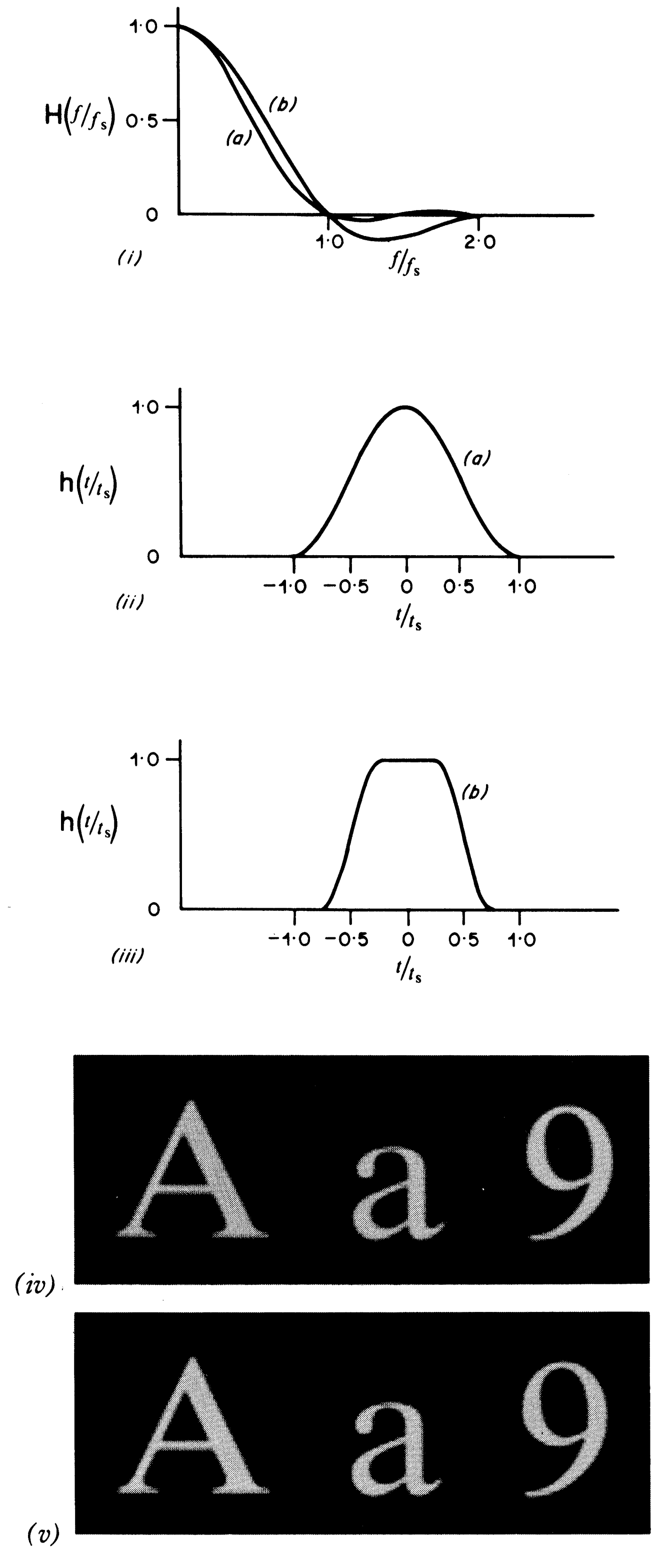


Fig. 16 — Raised-cosine time domain transition filters.

- (i) Frequency characteristics a and b
- (ii) Impulse response of filter a
- (iii) Impulse response of filter b
- (iv) Characters filtered using filter a
- (v) Characters filtered using filter b

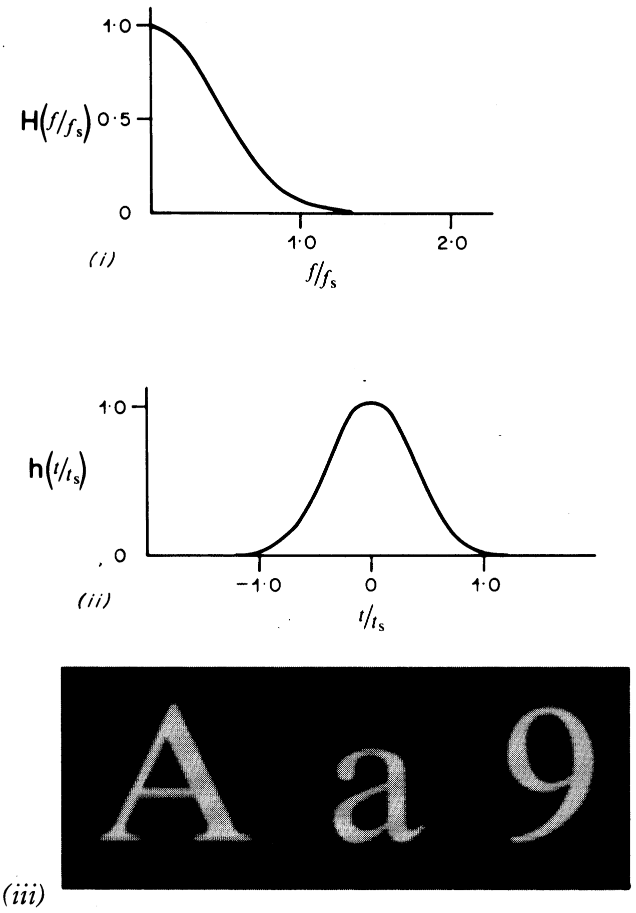


Fig. 17 — Gaussian filter.

- (i) Frequency characteristic of filter
- (ii) Impulse response of filter
- (ii) Characters filtered using this filter

with the Fourier transform of the windowing function.

Fig. 22 shows the impulse response and final frequency characteristic of the combination of a rectangular low-pass filter with raised-cosine transition band (as in (3) above and Fig. 21) and a Gaussian window function applied to the impulse response. The result is a pleasingly sharp letterform free of flicker and stepping but with fringeing restricted to the immediate vincity of each edge. The image is considerably sharper than that displayed with the 'camera-like' filter function of Section (1) and sharper than the image available from real scanned artwork. Similar results are obtained with 'raised-cosine' window functions.

5.3.4. Combined filtering

The use of separate horizontal and vertical filtering functions (a variables-separable approach) has been prescribed by the size and architecture

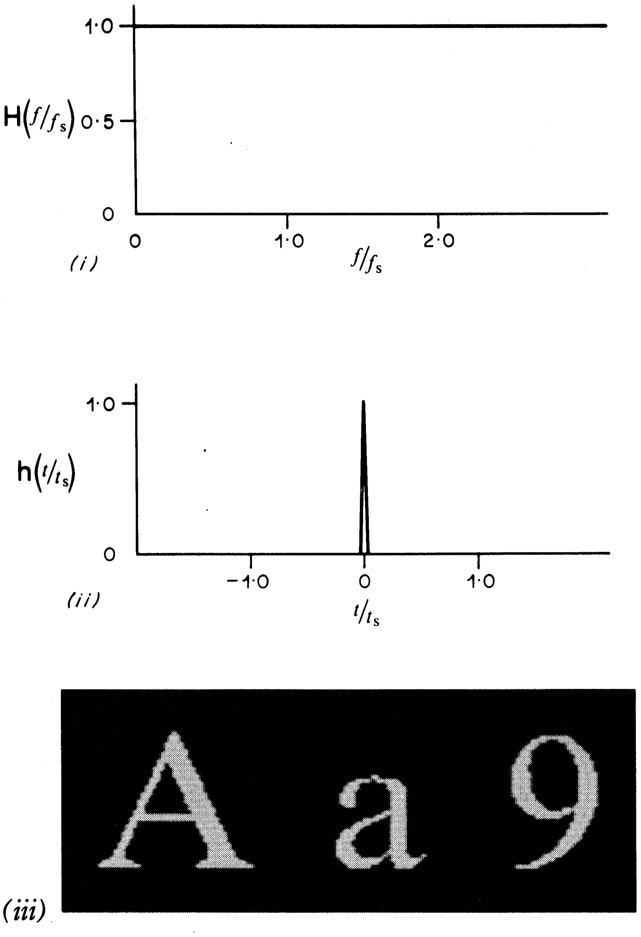


Fig. 18 — Undersampling.

(i) Frequency characteristic of undersampling
(ii) Effective 'impulse response' of undersampling
(iii) Characters undersampled showing characteristic stepping

of the computer and the size of the input character images (>1.2 Mbits as bit-map images). It is doubtful whether much improved results would be obtained from variables-non-separable filtering (combined horizontal and vertical filtering) except if the sampling structure is non-orthogonal (i.e. the horizontal sampling frequency is not line-locked).

It is likely that improved movement portrayal would be achievable for moving images (e.g. rolling captions) if some combined spatio-temporal filtering were to be performed (Fig. 13). This is beyond the scope of the equipement used and the spatial filtering techniques described here will certainly produce better motion portrayal than is available with unfiltered synthetic images. Further work will be necessary to produce combined spatio-temporal filtering which maintains the quality of text image which can be achieved on stationary images.

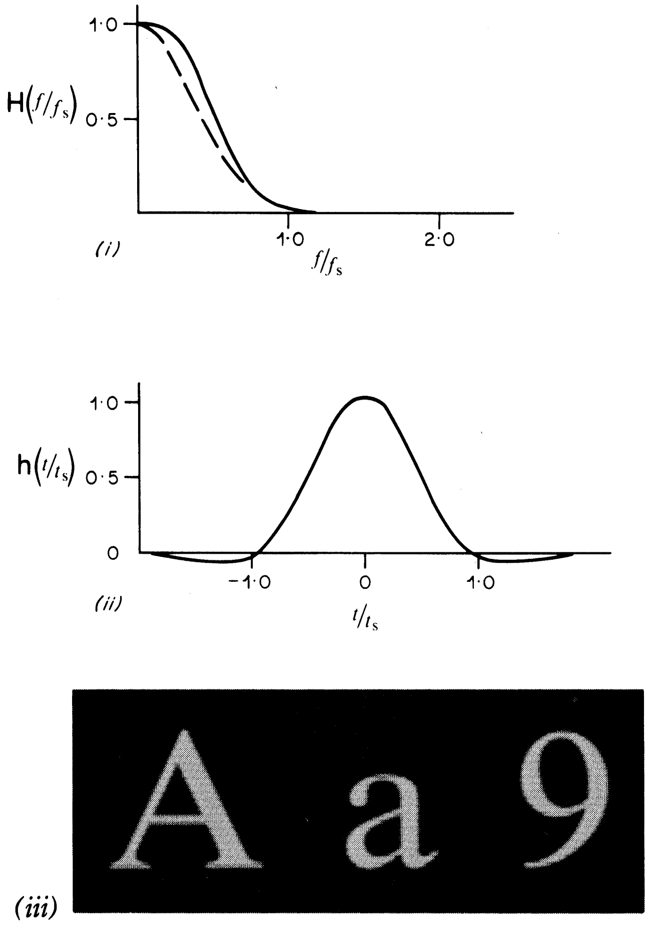


Fig. 19 — Aperture-corrected camera characteristic.

(i) Frequency characteristic — full line represents characteristic of corrected filter, dashed line that of uncorrected gaussian
(ii) Impulse response of 'corrected gaussian' filter
(iii) Characters filtered using 'corrected gaussian' filter

5.3.5. Number of bits required

By reducing the number of bits representing the display image signal transitions, under certain circumstances a respectable image can be displayed using 16 grey levels (4-bits). These are specifically when the filter function is monotonic (i.e. there are no overshoots to be represented) and when the image is linearly described as for soft-keying. Aliasing does become evident as the bits are further reduced and although no work has been done to confirm it, it is especially likely that movement portrayal will be degraded on rolling captions.

5.3.6. 'Pixel-fit'

If the master data has been properly filtered the perceived shape of the output letterform will be a perfect representation of the master artwork.

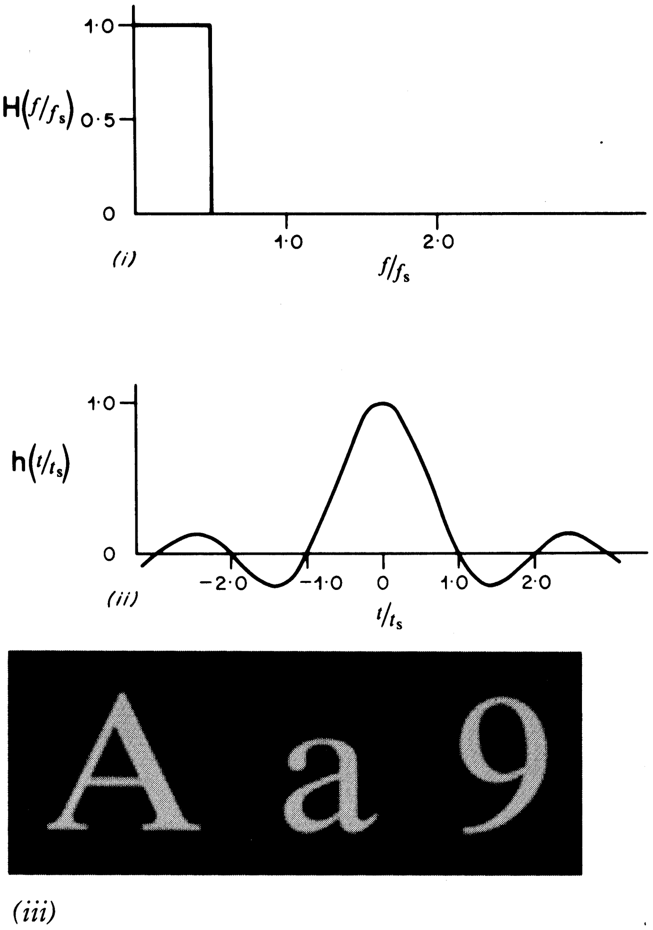


Fig. 20 — 'Ideal' low-pass filter.

(i) Frequency characteristic of filter
(ii) Impulse response of filter
(iii) Characters filtered with 'Ideal' impulse response showing extent of ringing around signal transitions

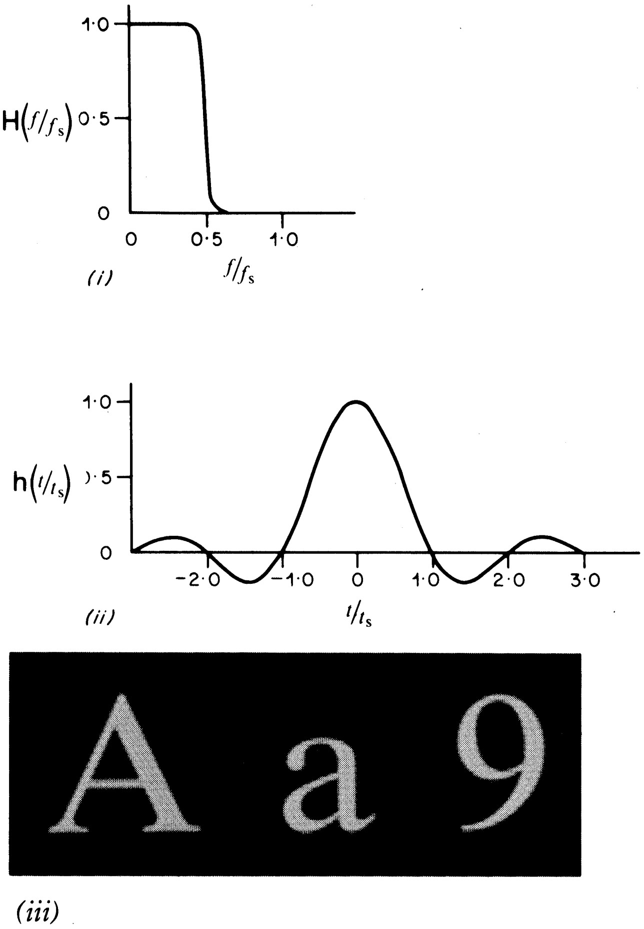


Fig. 21 — Less sharp-cut low-pass filter.

(i) Filter characteristic with 'raised-cosine' transition band
(ii) Impulse response of filter
(iii) Characters filtered using this filter

Unlike 2-level character images, every input sample in the master data makes a contribution to the final displayed image. There should be no need therefore to edit pixels or to adjust the sampling phase of the filtering function in order to achieve a better typographic representation of the master artwork. Residual aliasing, insufficient display grey levels or incorrect display gamma may, however, slightly corrupt the letterform and its proportions. The levels of aliasing remaining with the practical filters described in this Report are unlikely to affect the letterforms proportions but the other factors will do so.

5.3.7. Inter-character spacing

One of the complaints levelled at electronic character generators concerns the coarseness of control of inter-character spacing which is usually provided. Conventional artwork may be spaced with miniscule precision but character generators are usually limited to the horizontal pixel dimension or wider. By adjusting the exact phase of the output sample lattice with respect to the input sample lattice, filtered text may be adjusted in position and displayed to an accuracy determined by the input sample size.

For example, if the master data (2-level) has a horizontal resolution of 8 times the final output image then the output image may be positioned to within one eighth of a horizontal pixel dimension. Fig. 23 shows the reconstructed filtered edges corresponding to a step in master artwork 8 x the requisite output resolution. The edge has been filtered with a perfect low-pass filter (quite feasible using software) and the continuous waveform reconstructed using a practical (finite rate-of-cut) low-pass filter. The artwork has then been moved by one input pixel width and a new filtered continuous waveform constructed.

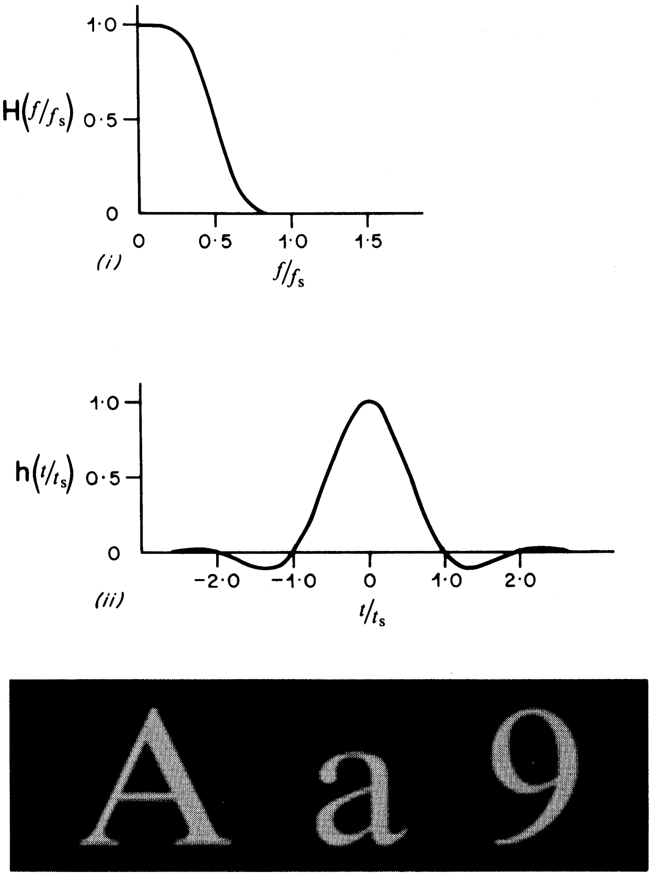


Fig. 22 – Windowed impulse response filter.

(i) Filter characteristic of a suitable filter
(ii) Impulse response of filter
(iii) Characters filtered using this filter

From this it can be seen that inter-character spacing can be adjusted with filtered images to a finer accuracy than the output pixel dimensions. In the case of Fig. 23 if the eight edges were all stored as separate character strokes and the horizontal output pixel resolution was 74 ns a spacing resolution of 9 ns would be possible. It is not necessary to use high horizontal sampling frequencies to achieve good positional accuracy or control.

6. Use of multi-level filtered text

Digitally-generated properly filtered text was used for the first time by the BBC for the General Election programme ELECTION '83 in June 1983. During this live programme, much use was made of electronic graphics images from computer controlled electronic painting systems. The text characters for these images were incorporated with other picture material as pixel data derived using the techniques described in this Report. The filtering functions used were of Gaussian form.

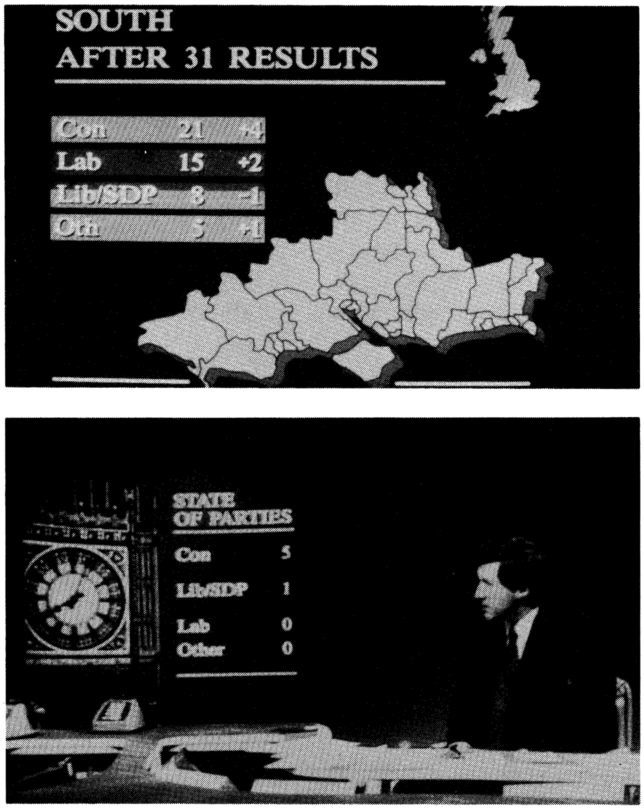


Fig. 24 – Examples of multi-level text used during the Election '83 broadcast in June 1983.

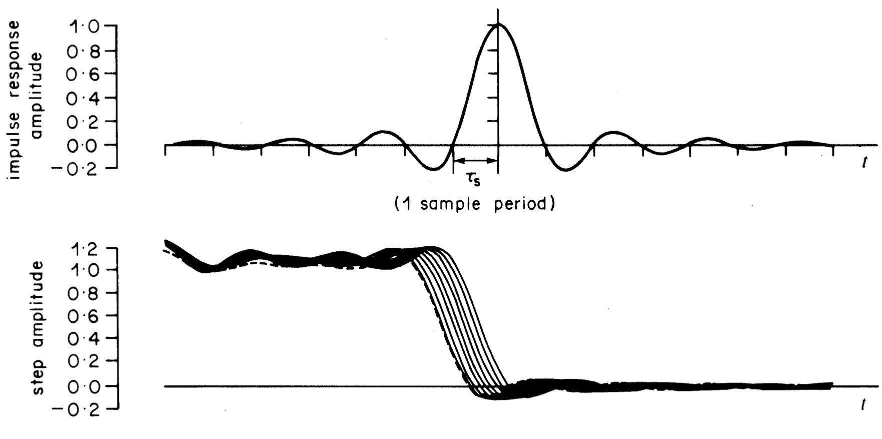


Fig. 23 – Sub-pixel adjustment of intercharacter spacing.

The continuous waveforms of the lower trace have been reconstructed from data sample values using a reconstructing low-pass filter whose impulse response is shown in the upper trace. The signal transition has been sampled at 8 different sampling phases corresponding to successive shifts of 1/8th sample period increments even though it is represented by data values at sample period increments.

Fig. 24 shows some examples of the use of multi-level text for the ELECTION '83 programme.

7. Conclusions

The image prefiltering performed by slide-scanners and television cameras serves to reduce the spectral components in the scanned picture signal which give rise to aliasing distortions when the picture is displayed. Electronically generated images (including electronic captions) are usually unfiltered and therefore display noticeable aliasing distortions in the form of 25 Hz flicker, stepping and poor movement portrayal. By carefully filtering the electronically generated images and representing them in a filtered form it is possible to achieve a displayed image quality comparable to, or better than, that of scanned artwork. In applying these techniques to electronic character generation it should be possible to make electronically generated titles and credits with all the freedom of typographic control and with the image quality of scanned artwork. The techniques are now applied to digital text generation in versions of the Quantel 'Paintbox' and Logica 'Flair' electronic graphics systems and in the Quantel 'Cypher' picture and text manipulator.

Antialiassed character generation will shortly become available in the Aston IV character generator.

8. Acknowledgements

The author would like to acknowledge the loan of master fount data by The Monotype Corporation and the advice and encouragement of a number of members of the BBC Graphic Design Department, in particular Charles McGhie and John Cook.

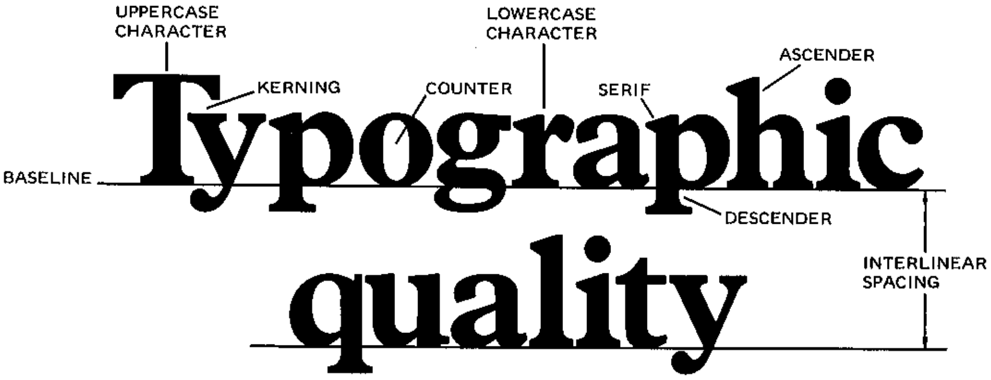
9. References

1. CLARKE, C.K.P. *and* TANTON, N.E. Digital standards conversion: interpolation theory and aperture synthesis. BBC Research Department Report, due for publication 1984.

2. WESTON, M. 1978. Digital Aperture Correction. BBC Research Department Report 1978/10.

10. Appendix

Glossary of typographic terms



ref A1 : CRAIG, H. 1978, Phototypsetting : A Design Manual. Pitman 1978 ISBN 0273 012800